Proceedings

ICSIIT 2010

International Conference on
Soft Computing, Intelligent System and Information Technology

1-2 July 2010

Bali, Indonesia

Editors:
Leo Willyanto Santoso
Andreas Handojo

Informatics Engineering Department
Petra Christian University

Center of Soft Computing and
Intelligent System Studies

Proceedings

International Conference on Soft Computing, Intelligent System and Information Technology 2010

Publisher:
Informatics Engineering Department
Petra Christian University

ISBN: 978-602-97124-0-7

Additional copies may be ordered from:
Informatics Engineering Department
Petra Christian University, Siwalankerto 121-131, Surabaya 60236, Indonesia

Cover Art production by Adi Wibowo/Informatics Engineering Department

Knowledge & Data Engineering

Development of Supporting Sales Analysis Application Using Frequent Closed Constraint Gradient Mining Algorithm (FCCGM)

Susana Limanto
Informatics Engineering
University of Surabaya
Raya Kalirungkut, Surabaya 60292 Indonesia
(031)2981395
susana@ubaya.ac.id

Dhiani Tresna Absari
Informatics Engineering
University of Surabaya
Raya Kalirungkut, Surabaya 60292 Indonesia
(031)2981395
dhiani@ubaya.ac.id

ABSTRACT

Making a promotion on certain items is a strategy to gain the competitive advantage in retail business. If a customer purchases item A for example, promotion on item A can be done by giving a bonus of item B, providing discount for item B, or any other form of promotions. At first glance, each kinds of promotion may decrease or even make harm the retail business. To avoid these conditions, the retail entrepreneur must be able to obtain proper information by mining their sales data. Information that can be gained is about set of items were purchased usually by customers in conjunction with particular items which appears to increase the average profit of that particular item. Based on this information, the sales volume of any particular item may be raised by making the particular item as a promotion item that must be sold together with certain set of items.

The most common obstacles in getting information from sales data is the limited human ability to process large amounts of data efficiently. This research is proposed to help entrepreneurs to simulate the retail sales data in generating information about the set of items that frequently purchased with a particular item which will increase the average profit of a particular item, using the FCCGM algorithm.

Keywords

data mining, frequent closed constrained gradient mining, sales analysis.

1. INTRODUCTION

Recently, retail business has been became a business that attract many entrepreneur attentions. This can be proved by the increasing number of existing retail businesses at the nearby locations, such as Alfa Mart, Alfa Midi, Indomaret and Circle K. The number of emerging retail business will increase business competition. So, many promotions are made to attract as many customers to come to their stores.

Good promotion should not harm the respective retail businesses and could attract customers. Unfortunately, not all promotion attracts customers. Retail entrepreneur must be able to establish a strategy based on the customer's shopping habits to avoid condition as mentioned before. Information about customer's shopping habits can be gained from sales data and can be used to get information about the set of items that were usually purchased in conjunction with certain items. The set of items that are sold together with some promotion items might increase or decrease the average profit of those promotion items. *Frequent Closed Constrained Gradient Mining* (FCCGM) is an algorithm which can be used to obtain information about the set of items that will increase the average profit of promotional items.

2. FREQUENT CLOSED CONSTRAINED GRADIENT MINING (FCCGM) ALGORITHM

FCCGM algorithm can be used to determine set of items that were usually purchased by customers in conjunction with any promotional item, which may increase average profits of promotion item as desired. Those set of items is known as frequent closed constrained gradient item sets. This algorithm was first introduced by Wang et al. in 2006 and was implemented on sales data (also known as transactional data) in a retail database. This algorithm comprises six main processes. The processes are calculation of measure value, followed by construction of projected database, deletion of the items that do not meet *the minimum* criteria for *support* and *gradient threshold,* construction of FP-Tree, deletion of items that do not satisfy the gradient threshold and the last process is to mine remaining data. All of these processes will be explained below.

The first process is calculating the *measure* value. *Measure* is defined as average profit of promotion items in each transaction. After the calculation of measure value, projected database is constructed. Projected database is constructed by excluding the transactions without promotion items, removing the promotion items from remaining transactions, and then sorting remaining items in all transactions according to support value in descending order. Support value is defined as the frequency of appearance a group of specific items in all existing transactions **(Megaputer, 2000).**

The third process consists of few steps. The first step is to remove the items that do not meet the *minimum support* from *the projected database* and to calculate *the gradient threshold* value. *Gradient threshold* value is defined as sum of all transaction *measure value* in *the projected database,* divide by the number of transactions and time by defined *minimum gradient.* The next step is to calculate the top-*K average value* of each item. *Top-K average* value of an item Z is average *measure* of the first K